NUMERICAL ANALYSIS PROJECT
MANUSCRIPT NA-92-12

SEPTEMBER 1992

Matrix Shapes Invariant under the Symmetric QR Algorithm

by

Peter Arbenz
and
Gene H. Golub

NUMERICAL ANALYSIS PROJECT
COMPUTER SCIENCE DEPARTMENT
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

MATRIX SHAPES INVARIANT UNDER THE SYMMETRIC QR ALGORITHM *

PETER ARBENZ
Eidgenössische Technische Hochschule
Institut für **Wissenachaftliches** *Rechnen*
CH-8092 **Ziirich, Switzerland**
arbenz@inf.ethz.ch

GENE H. GOLUB
Stanford **University**
Department of Computer **Science**
Stanford CA 94305-2140, U.S.A.
golub@cholesky.stanford.edu

ABSTRACT
It is shown, which zero patterns of symmetric matrices are preserved under the QR algorithm.

Keywords: QR algorithm, zero pattern.

1. Introduction

The QR algorithm

$$\begin{aligned} A_{k-1} - \sigma_k I &=: Q_k R_k \qquad \text{(QR decomposition)} \\ A_k &:= R_k Q_k + \sigma_k I \end{aligned} \tag{1}$$

is the standard algorithm for computing the spectral decomposition of a symmetric matrix $A = A_0$ [6, §§8-14], [7, §4], [2, §2.3.3]. The algorithm computes a sequence $\{A_k\}_{k=1}^{\infty}$ of symmetric matrices similar to A converging to a diagonal matrix. Appropriately chosen shifts σ_k increase the rate of convergence from linear to cubic [6].

As a QR iteration step costs $O(n^3)$ floating point operations if it is applied to a dense matrix of order n, the algorithm is practically used only after the original

• **We dedicate this paper to Velvel Kahan and Beresford Parlett who have led the way in our understanding of the fundamentals of matrix computations.**

matrix A has been transformed to a similar matrix of simpler form, i.e. with much fewer nonzero elements.

It is well-known that the symmetric QR algorithm (1) preserves the banded structure. It is therefore most often applied to tridiagonal matrices, which reduces the complexity of one QR step to $O(n^2)$ or even $O(n)$ if only the eigenvalues of A are wanted.

Until now it has remained an open question as to which matrix structures are preserved using the QR method described by (1). Recently, Arbenz and Golub [3] have shown that the QR algorithm does not preserve arrow matrices, i.e. matrices whose elements vanish except those on the diagonal and in the first row and column.

In this note we answer the general question which zero structures *are* preserved by the symmetric QR algorithm (1).

2. Statement of results

In order to state our results we make two definitions.

A matrix A is said to be ***reducible,*** if there exists a permutation matrix P such that

$$PAP^T = \begin{pmatrix} A_1 & A_{12} \\ O & A_2 \end{pmatrix}.$$

Otherwise, it is called ***irreducible.*** If A is symmetric, A_{12} is zero, of course.

A **staircase** *matrix* is a matrix for which $m_j(A) \geq m_{j-1}(A)$, $j > 1$, where

$$m_j(A) := \max\{j, \max_{i>j}\{i | a_{ij} \neq 0\}\}$$

is the index of the last nonzero element of the j-th column of A. Two staircase

$$A_1 = \begin{pmatrix} \times & \times & & & \\ \times & \times & \times & \times & \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \end{pmatrix} \qquad A_2 = \begin{pmatrix} \times & \times & & & \\ \times & \times & & \times & \\ & & \times & & \times \\ & \times & & \times & \times \\ & & \times & \times & \times \end{pmatrix}$$

FIG. 1. *Examples of staircase* **matrices with** $m_1 = 2, m_2 = 4, m_3 = m_4 = m_5 = 5$.

matrices with the same parameters m_j are depicted in Figure 1. (As usual, blank areas denote zero elements.) To emphasize that there are no zero elements within the stair, we denote the matrix A_1 on the left hand side to be *full* staircase.

The following Theorem identifies the matrix patterns preserved under the symmetric QR algorithm.

Theorem 1. Let A be a symmetric matrix.

1. If A is reducible through the permutation matrix P then all QR iterates $A^{(k)}$ are reducible by means of the same P.

2. If A is irreducible then the zero pattern of A is preserved by the QR algorithm if and only if A is a full staircase matrix.
   There is fill-in in all other situations: After a number of iterations the QR iterates will be full staircase matrices, the staircase being the smallest that contains the nonzero structure of A.

This means that to check whether a certain zero pattern of a matrix A *is* preserved, one first has to permute A into a direct sum of irreducible submatrices and then check whether these are full staircase. The determination of irreducible submatrices of a given matrix or, equivalently, of a connected subgraph of a graph is a well-known problem in computer science and is solved by standard algorithms (see e.g. [1, §7]).

Notice that we consider only structures of matrices, here. There are, of course, instances of initial matrices $A = A^{(0)}$ that lead (by numerical cancellation) to iterates $A^{(k)}$ that have zeros within the stair (cf. [5, p. 27]).

If A is nonsymmetric, statement 1 in Theorem 1 does not hold. This is immediately seen when applying one QR iteration step to a matrix of the form

$$A = A^{(0)} = \begin{pmatrix} x & x & x & x \\ & x & & x \\ x & x & x & x \\ & x & & x \end{pmatrix} .$$

In this example, the elements $a_{2,1}^{(1)}$, $a_{2,3}^{(1)}$, and $a_{4,3}^{(1)}$ of $A^{(1)}$ are non-zero. The staircase form is however a sufficient condition for shape preservation for nonsymmetric matrices. Notice that this means that the upper triangle of a nonsymmetric matrix fills up. This is even true if A_0 is a nonsymmetric tridiagonal matrix which is similar to a symmetric tridiagonal matrix.

An interesting matrix structure which is preserved by the QR algorithm because of part one of the Theorem is a bandmatrix of bandwidth b, whose off-diagonal elements vanish except those on the extremal b-th (cf. Fig. 2).

$$A = \begin{pmatrix} x & & x & & \\ & x & & x & \\ x & & x & & x \\ & x & & x & \\ & & x & & x \end{pmatrix}$$

FIG. 2. *Example of an almost empty bandmatrix ($n = 5$, $b = 2$).*

3. Proof of Theorem 1

We first show that reducibility is preserved by the symmetric QR algorithm.

Lemma 2. Let A be reducible. Then there is a permutation matrix P such that $PAP^T = A_1 \oplus A_2$ and

$$PA'P^T = A_1' \oplus A_2',$$

where A_1', A_2', respectively, are obtained from A, A_1, A_2, by one QR step with the same shift.

Proof. Without loss of generality we can set the shift to zero.

Let π be a permutation of the set $\{1, \ldots, n\} \subset \mathbb{N}$. Then we define the permutation matrix P_π by $P_\pi \mathbf{e}_{\pi(k)} = \mathbf{e}_k$. If P_π acts on a vector, the element with index $\pi(k)$ is moved to position k. We now choose π such that

$$P_\pi A P_\pi^T = A_1 \oplus A_2, \qquad A_1 \in \mathbb{R}^{k \times k}, \quad A_2 \in \mathbb{R}^{(n-k)\times(n-k)},$$

and

$$\pi(1) < \cdots < \pi(k), \quad \pi(k+1) < \cdots < \pi(n),$$

i.e. within A_1 and A_2, rows and columns appear in their original order.

Let $A_j = Q_j R_j$ be a QR factorization of A_j, $j = 1, 2$. Then

$$A = P_\pi^T (A_1 \oplus A_2) P_\pi = P_\pi^T (Q_1 \oplus Q_2) P_\pi P_\pi^T (R_1 \oplus R_2) P_\pi^T.$$

$R := P_\pi (R_1 \oplus R_2) P_\pi^T$ is upper triangular. In fact, let $j > i$. Then

$$r_{ji} = \mathbf{e}_j^T P_\pi (R_1 \oplus R_2) P_\pi^T \mathbf{e}_i = \mathbf{e}_{\pi^{-1}(j)}^T (R_1 \oplus R_2) \mathbf{e}_{\pi^{-1}(i)}.$$

$\pi^{-1}(k)$ is that index, which becomes k after application of P_π. If $\pi^{-1}(j) > \pi^{-1}(i)$, r_{ji} evidently vanishes. If $\pi^{-1}(j) < \pi^{-1}(i)$, then P_π changes the order of row i and j. Therefore, $\pi^{-1}(j) \le k < \pi^{-1}(i)$, and $r_{ji} = 0$ as it originates from the the upper-right zero part of $R_1 \oplus R_2$. Let $Q := P_\pi^T (Q_1 \oplus Q_2) P_\pi$. Then $A = QR$ is a QR factorization of A. $A' = Q^T A Q$ is obtained by one QR step applied to A. Then

$$\begin{aligned} P_\pi A' P_\pi^T &= P_\pi Q^T A Q P_\pi^T = (P_\pi Q^T P_\pi^T)(P_\pi A P_\pi^T)(P_\pi Q P_\pi^T) \\ &= (Q_1^T \oplus Q_2^T)(A_1 \oplus A_2)(Q_1 \oplus Q_2) = A_1' \oplus A_2', \end{aligned}$$

which is the claimed result. $\square$

By recursion, Lemma 2 applies to the case where

$$PAP^T = \bigoplus_{j=1}^{p} A_j, \qquad A_j \text{ irreducible},$$

for arbitrary p. Thus, part 1 of Theorem 1 is proved. The Lemma does not hold for nonsymmetric matrices as we noted at the end of §2.

Next we prove the evident

Lemma 3. Let $1 \le i < j \le n$. Then

$$a_{\iota\eta} = 0, \quad \iota \ge j, \eta \le i \qquad \Longrightarrow \qquad a'_{\iota\eta} = 0, \quad \iota \ge j, \eta \le i.$$

Proof. As $Q\mathbf{e}_\eta \in \operatorname{span}\{A\mathbf{e}_1, \ldots, A\mathbf{e}_\eta\}$ we have

$$q_{\iota\eta} = 0, \qquad \iota \ge j, \eta \le i.$$

But

$$a'_{\iota\eta} = \sum_{k=1}^{n} r_{\iota k} q_{k\eta} = 0,$$

as $r_{\iota k} = 0$ for $k < \iota$ and $q_{k\eta} = 0$ for $k \geq \eta$. □

This Lemma shows that the staircase shape is preserved by the QR algorithm. We did not make use of the symmetry of A whence Lemma 3 also holds for non-symmetric matrices.

We now show that the QR iterates A_k *become* full staircase after a certain number of iteration steps. This statement holds even if we restrict ourselves to the QR algorithm without shift applied to positive definite matrices.

It is well-known that one QR iteration step applied to positive definite matrices is equivalent to two Cholesky-LR iteration steps [7, p.321]. Therefore, we can make use of the known results concerning the fill-in when computing the Cholesky decomposition of large sparse matrices [5].

To that end, let $G^A := (X^A, E^A)$ be the adjacency graph of the matrix $A \in \mathbb{R}^{n\times n}$ with node set $X^A := \{1, \ldots, n\}$ and edge set $E^A := \{\{i, j\} \subset X^A \mid a_{i,j} \neq 0, i > j\}$. In an analogous way we define the adjacency graph $G^L := (X^L, E^L)$ of the Cholesky factor L of $A = LL^T$.

From symbolic factorization of large sparse positive definite matrices the following rule is well-known [5, p. 98]

Rule 1. $\{i, j\} \in E^L \iff$ There is a path $(i, k_1, \ldots, k_t, j)$ in E^A with
$1 \leq k_l < \min\{i, j\}$, $1 \leq l \leq t$.

Notice that t may be zero, i.e. that $\{i, j\} \in E^A$ already. So, $G^L \supset G^A$.

On the other hand it is immediately clear from $A' = L^T L$ that

Rule *2.* $\{i, j\} \in E^{A'} \iff \{i, j\} \in E^L$ or there is a $k > \max\{i, j\}$ with
$\{k, i\} \in E^L$ and $\{k, j\} \in E^L$.

In Figure 3 examples are given of matrices which are filled in a single QR iteration step. Left pointing arrow matrices are filled because of rule 1, right pointing arrow matrices because of rule 2 [4].

$$A_1 = \begin{pmatrix} \times & \times & \cdots & \times \\ \times & \times & & \\ \vdots & & \ddots & \\ \times & & & \times \end{pmatrix} \qquad A_2 = \begin{pmatrix} \times & & & \times \\ & \ddots & & \\ & & \times & \times \\ \times & \cdots & \times & \times \end{pmatrix}$$

Fig. 3. *Left* **pointing** (A_1) *and right* **pointing** (A_2) **arrow** *matrices.*

From the two rules above we see that $E^{A^{(k)}} \supset E^{A^{(k-1)}}$, $k > 0$. We now want to determine at what positions fill-in (if any) occurs in the course of the QR iteration. As we are not interested in the iteration step at which a particular fill-in actually happens we do not care about the real sequence with which rules 1 and 2 are applied in the algorithm. In contrast to what happens in the real algorithm, we apply the rules only once per iteration. So, in our analysis, we are lagging behind the real zero structure. However, as we will end up with a nonzero pattern, the full staircase form, which cannot by means of Lemma 3 be exceeded by the real algorithm, the final nonzero pattern will coincide with the desired one.

Proof of Theorem 1. First, we show that all zero elements $a_{i,1}^{(k)}$, $i < m_1$, fill in for some k. *So,* let $a_{i,1}^{(0)} = 0$, $1 < i < m_1$. As A is irreducible, there is a path ($k = x_0, \ldots, x_t = i$) in E^A from k to i. This path may or may not go through 1. By repeated application of rule 1 when $x_{j-1} > x_j < x_{j+1}$ and rule 2 when $x_{j-1} < x_j > x_{j+1}$ respectively, we get a path $(k = x'_0, \ldots, x'_{t'} = i)$ in $E^{A^{(k')}}$ for some k' with $x'_0 > x'_1 > \ldots > x'_{t'}$. Rule 2 can now be applied t' times to yield $\{i, 1\} \in E^{A^{(k'+t')}}$.

Remark. If $i > k$, an analogous procedure leads to a path $(k = x''_0, \ldots, x''_{t''} = i)$ in $E^{A^{(k'')}}$ for some k'' with $x''_0 < x''_1 < \cdots < x''_{t''}$. Neither of the two rules are now applicable. Therefore, $m_1(A^{(k)}) = m_1(A^{(0)})$ $k > 0$, which is in accordance with Lemma 3.

We have shown that $(i, 1) \in E^{A^{(k)}}$ for some k for $1 < i \leq m_1$. Rule 1 implies that $\{i, j\} \in E^{A^{(k+1)}}$ for $1 \leq i < j \leq m_1$. That means that all elements of the principal $m_1 \times m_1$ block of $A^{(k+1)}$ are nonzero. As there is fill-in, certain $m_j(A^{(k+1)})$, $j > 1$, may be larger than $m_j(A^{(0)})$. Rules 1 and 2 (and Lemma 3) however imply that

$$m_j(A^{(k+1)}) \leq \max_{1 \leq i \leq j} m_i(A^{(0)}).$$

Therefore, by restricting ourselves to the section graph of $G^{A^{(k+1)}}$ corresponding to the node subset (2, . . . , n} we show that there is a m such that $a_{i,2}^{(m)} \neq 0$, $2 \leq i \leq m_2$. Note that this section graph is connected.

Proceeding in this way with the node sets $\{i, \ldots, n\}$, $i = 3, \ldots, n$, *we* get the desired result. □

Acknowledgments.

The research of the second author was in part supported by NSF grant CCR-8821078.

References.

[1] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *Data Structures and Algorithms*, Addison-Wesley, Reading, MA, 1983.

[2] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. D. Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[3] P. Arbenz and G. H. Golub, *QR-like algorithms for symmetric arrow matrices*, SIAM J. Matrix Anal. Appl., 13 (1992), 655-658.

[4] J. Demmel and W. Gragg, *On computing accurate singular values and eigenvalues of acyclic matrices*, IMA preprint 962, Institute for Mathematics and its Applications, University of Minnesota, Minneapolis MN, April 1992. To appear in Linear Algebra Appl.

[5] A. George and J. W. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

[6] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, NJ, 1980.

[7] D. S. Watkins, *Fundamentals of Matrix Computations*, Wiley, New York, 1991.